JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, D15201, doi:10.1029/2009JD012041, 2009

Note on the weekly cycle of storm heights over the southeast United States

Thomas L. Bell,[1] Jung-Moon Yoo,[2] and Myong-In Lee[3]

Received 12 March 2009; revised 18 May 2009; accepted 3 June 2009; published 13 August 2009.

[1] An earlier paper by Bell et al. (2008) showed satellite evidence that average summertime (1998–2005) rainfall over the noncoastal southeast U.S. varied with the day of the week in a statistically significant way, with the maximum occurring midweek (Tuesday–Thursday). An explanation was proposed in which the recurring midweek increase in air pollution over the area causes a shift in the drop size distribution in clouds to smaller sizes as the clouds develop. The smaller droplets could be carried to higher altitudes where their freezing releases additional latent heat, invigorating the storms. Evidence for this phenomenon was provided by storm height distributions obtained from the Tropical Rainfall Measuring Mission radar, but the statistical significance of the midweek increase in storm heights was unclear. An improved statistical analysis of the storm height distributions is provided here, indicating that the probability that storms climb above altitudes of 7–15 km is increased midweek relative to weekends (Saturday–Monday) for afternoon storms (1200–2400 LT). The morning storm heights, on the other hand, are found not to exhibit statistically significant shifts, which would be consistent with the above explanation. Morning storm statistics are also found to be much more sensitive than afternoon storm statistics to the exact area over which the averages are taken.

Citation: Bell, T. L., J.-M. Yoo, and M.-I. Lee (2009), Note on the weekly cycle of storm heights over the southeast United States, *J. Geophys. Res.*, *114*, D15201, doi:10.1029/2009JD012041.

1. Introduction

[2] In an earlier paper by *Bell et al.* [2008] (hereinafter referred to as B08) evidence was presented for a weekly cycle in rain rate estimates from the Tropical Rainfall Measuring Mission (TRMM) satellite's microwave instruments over the noncoastal southeast U.S. This area was referred to by B08 as "area B" and is shown in Figure 1 as the red crosshatched area. Averages of TRMM rain estimates over area B for the summertime (June–August) for 1998–2005 showed substantial changes in average rainfall with the day of the week, peaking on Tuesday and remaining large for the next 2 days.

[3] The explanation of the dependence of rain rate on the day of the week proposed by B08 invoked the well-known variations in pollution with the day of the week and the theory described by Rosenfeld in the papers by *Williams et al.* [2002] and *Andreae et al.* [2004] and further developed by *Rosenfeld et al.* [2008] that the decrease in droplet sizes in storm clouds forming in "dirty" air enabled more liquid water to reach higher altitudes and to release additional latent heat as it froze, energizing the storms and causing them to grow larger and rain more. The mechanism requires that the storms form in environments such as those that exist in the southeast United States during the summertime: highly unstable vertical temperature structures, with ample moisture below and cloud base temperatures well above freezing. This theory would not apply to the drier western half of the country, and indeed, no weekly cycle was discernible there. The theory predicts, in fact, that this effect of pollution should be maximum in the afternoons, and this was observed: the statistical significance of the weekly cycle in rainfall over area B increased considerably when averages were restricted to afternoon (1200–2400 LT) data.

[4] This evidence was reinforced by B08's analysis of surface rain gauge data and the model reanalysis data (version R-2 of the National Centers for Environmental Prediction and the Department of Energy reanalysis data) [*Kanamitsu et al.*, 2002]. As the satellite data suggested, daily rainfall as measured by rain gauges increased in the middle of the week, and lower-level wind convergence, upper-level divergence, and 500-hPa vertical winds over area B changed with the day of the week in a way that was consistent with the changes in convection implied by the rain activity. Furthermore, although not reported by B08, the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard both NASA's Terra and Aqua satellites [*Remer et al.*, 2005] also shows significant increases in fractional cloud cover during the middle of the week over area B, accompanied by decreases in cloud top temperatures, both signs of increased midweek convective activity.

[1] Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[2] Department of Science Education, Ewha Womans University, Seoul, South Korea.
[3] Goddard Earth Sciences and Technology Center, University of Maryland Baltimore County, Baltimore, Maryland, USA.

0148-0227/09/2009JD012041$09.00

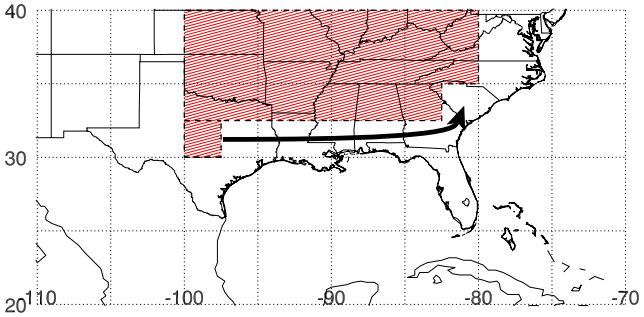

Figure 1. Averaging area (red cross-hatching) used by *Bell et al.* [2008], called "area B" therein. A new rectangular averaging area is used here, by moving the grid box at the bottom to the right-hand side, as indicated by the arrow. This rectangular area is referred to here as "area B′."

[5] The TRMM satellite precipitation radar (PR) provided additional evidence of storm invigoration, showing that the distribution of storm heights shifts to higher altitudes during the middle of the week compared to weekends. At the time of B08's publication a credible statistical analysis of the changes in the PR storm height distributions with the day of the week was unavailable. This note is intended, in part, to rectify that. In the following paper we describe the PR storm height data used in the analysis, present the method of estimating the statistical confidence of the changes we see, provide some discussion of how sensitive the changes in the morning (0000–1200 LT) distributions are to averaging details, and offer our conclusions.

2. Description of Data

[6] The TRMM PR product 2A23 (TRMM PR Team, Tropical Rainfall Measuring Mission (TRMM) precipitation radar algorithm: Instruction manual for version 6, 2005, available at http://www.eorc.jaxa.jp/TRMM/document/pr_manual/pr_manual_v6.pdf) reports storm height for each radar observation where the PR algorithm (version 6) determines that precipitation is detected within the radar beam with a high degree of confidence. "Storm height" here means the height of the highest point in the radar beam with detectable returns (∼17–18 db$Z$), measured relative to mean sea level. The PR footprint is roughly 4–5 km in diameter. A more complete discussion of the issues involved with the interpretation of this product was provided by B08. We analyze storm height data here for the same period used by B08, 1998–2005 summers (June–August).

[7] It should be noted that because of the TRMM's low-inclination orbit and the PR's swath width, the PR is unable to see north of about 36.3°. The PR's observations are most frequent in the neighborhood of latitude 33.7°N, and our areal statistics consequently weight the higher latitudes more, proportional to the PR's observational frequency.

3. Method of Analysis

3.1. Statistical Measures Used

[8] Instead of averaging over the irregularly shaped area B in Figure 1, we used a simpler, rectangular box spanning latitudes 32.5–40°N and longitudes 100–80°W. This is the area formed by substituting for the southernmost 2.5° × 2.5° grid box of area B the bottom right-hand corner of the rectangle, as indicated by the arrow in Figure 1. We shall refer to this area as "area B′."

[9] Histograms $n(a)$ of storm heights in area B′ are obtained by counting the number of PR footprints identified by the PR algorithm as containing rain and with storm heights in an altitude bin labeled by altitude a (in km), where the bin extends from $a - 0.5$ km to $a + 0.5$ km. These histograms are used to calculate the fraction of the footprints in area B′ for which storm heights are in bin a or above for local observation times falling on either Tuesday–Thursday (TWT) or Saturday–Monday (SSM). We further subdivide the observations into morning (0000–1200 LT) and afternoon (1200–2400 LT) categories. If, for instance, $n_{\mathrm{TWT,m}}(a)$ is the number of footprints in area B′ for Tuesday–Thursday mornings with storm heights in bin a, then we can write the fraction of footprints with storm heights at a or above as

$$c_{\mathrm{TWT,m}}(a) = \frac{\sum_{a'=a}^{t} n_{\mathrm{TWT,m}}(a')}{\sum_{a'=0}^{t} n_{\mathrm{TWT,m}}(a')}, \tag{1}$$

where the denominator in (1) is in effect just the total number of footprints with PR-detected rain and t is the bin with maximum reported storm height. Because no storm heights above 20.5 km are detected, $t = 20$.

[10] As a measure of the change in height distributions with the day of the week, we investigate the ratios

$$r_i(a) = c_{\mathrm{TWT},i}(a)/c_{\mathrm{SSM},i}(a), \tag{2}$$

where the index $i = \{\mathrm{m}, \mathrm{a}\}$ indicates whether the data are for mornings or afternoons. The ratio $r_i(a)$ tells us how much more probable it is that storms reach or exceed altitude a Tuesday–Thursday (midweek) compared to Saturday–Monday (weekends). If there were no change in behavior with the day of the week, we would expect $r_i(a) = 1$.

3.2. Sampling Error Estimates

[11] The sampling error in the ratio is represented by δr_i:

$$r_i = \langle r_i \rangle + \delta r_i \, . \tag{3}$$

(We omit specifying the altitudes a here and in equation (4) to help simplify the notation.) The angle brackets in (3) indicate the expected values of the quantities that we would calculate if we had infinite amounts of data with the same climatological statistics as the data we actually have. Our best estimate of these expectations will in fact be the averages obtained from the data we actually have. Note that, by definition, $\langle \delta r_i \rangle = 0$.

[12] B08 (Appendix A, equation (A8)) give an estimate of the sampling error variance in these ratios,

$$\mathrm{Var}(\delta r_i) \approx \langle c_{\mathrm{SSM},i} \rangle^{-2} \mathrm{Var}\left(\delta c_{\mathrm{TWT},i}\right) + \langle c_{\mathrm{TWT},i} \rangle^{2} \langle c_{\mathrm{SSM},i} \rangle^{-4} \cdot \mathrm{Var}\left(\delta c_{\mathrm{SSM},i}\right), \tag{4}$$

which we have copied above, but with the dependence on whether it is morning or afternoon, i (either m or a), made explicit.

[13] The variances $\mathrm{Var}(\delta c_{\mathrm{TWT},i})$ and $\mathrm{Var}(\delta c_{\mathrm{SSM},i})$ needed in (4) are estimates of the sampling error variances in the individual distributions $c_{\mathrm{TWT},i}$ and $c_{\mathrm{SSM},i}$. B08 estimated these error variances by assuming that storm heights for each PR footprint are statistically independent of each other. The estimates ignored the effects of spatial and temporal correlations in the data, which are surely substantial, with the result that the error estimates by B08 were at best lower bounds for the actual errors. We improve on these error estimates here so that we get a better sense about which changes in the storm height distributions are "real."

[14] To try to deal with spatial and temporal correlations in the storm heights, we work with the storm height distributions for each week, represented by $n_{p,w}(a)$, where p denotes the period from which the data come: whether the data are from TWT or SSM and whether they are for mornings or afternoons (i is m or a). Weeks are labeled by integers $w = 1, \ldots, W$. Thus,

$$n_{\mathrm{TWT,m}}(a) = \sum_{w=1}^{W} n_{p,w}(a), \; p = \{\mathrm{TWT}, \mathrm{m}\}, \tag{5}$$

where the sum is over all W weeks during the summers of the years of interest. There are 13 weeks per summer. Storm behavior is scarcely predictable beyond a few hours, and it is reasonable to assume that the distributions $n_{p,w}(a)$ are not very correlated from week to week. By treating the weekly distributions as a single "measurement," most of the statistical effects of spatial and temporal correlations are captured.

[15] We assume that the sampling error variance of an average over W uncorrelated observations x_w can be approximated when W is large by the normal-statistics result

$$\mathrm{Var}(\overline{x_w}) = \frac{1}{W}\mathrm{Var}(x_w) \; , \tag{6}$$

where the bar notation indicates the average

$$\overline{x_w} = \frac{1}{W}\sum_{w=1}^{W} x_w \tag{7}$$

and Var is the estimated variance of the variable x_w

$$\mathrm{Var}(x_w) = \frac{1}{W-1}\sum_{w=1}^{W} (x_w - \overline{x_w})^2. \tag{8}$$

This gives us estimates of the error variance in the overall storm height distributions in equation (5) for large W:

$$\mathrm{Var}\left[\delta n_p(a)\right] \approx W\mathrm{Var}\left[n_{p,w}(a)\right], \tag{9}$$

where $\mathrm{Var}[n_{p,w}(a)]$ is estimated as in equation (8). (Note that the familiar factor $1/W$ is replaced by W on the right-hand side of equation (9) because the factor $1/W$ is absent from the definition of $n_{\mathrm{TWT,m}}(a)$ in equation (5).)

[16] Building on this estimate for the error variance of $\delta n_p(a)$, we can make estimates of the error variances of the ratios r_i since they are all derived from the distributions $n_p(a)$ through equation (1). This is the basis for estimating the sampling error of $r_i(a)$ in equation (4). Details are given in Appendix A.

4. Storm Height Distributions

[17] As in the work by B08, PR data for storm heights for the summers of 1998–2005 were analyzed, but over area B′ and with error bars for the ratios $r_i(a)$ estimated as described in Appendix A. We verified that the number of PR observations was about equally distributed over all hours and all days of the week during the course of the eight summers, as might be expected, since the TRMM satellite local observation time progresses through all 24 h during the course of about 46 days, or about half of one summer [*Negri et al.*, 2002]. The results for the ratios $r_i(a)$ are shown in Figure 2b with 1-sigma error bars. Figure 2b should be compared to Figure 10c in the paper by B08, which is reproduced here in Figure 2a. As was found by B08, storms in the afternoon tend to climb to higher altitudes during the midweek period compared with weekends, particularly for those reaching altitudes above 7–15 km. About 20% of afternoon storm heights reach the 9-km bin or higher on weekends, and 40% more midweek storm heights exceed 9 km than weekend storm heights. As expected, the error bars found here are generally much larger than those estimated by B08, almost certainly because of the amount of spatial and temporal correlation in the data.

[18] The behavior of the morning data shown in Figure 2 is, however, somewhat different from what was found in the work by B08, where the statistics for $r_{\mathrm{m}}(a)$ appeared to show that the ratios for morning storm heights were significantly above 1 at higher altitudes. Our new results are consistent with there being no change with the day of the week in morning storm height distributions.

[19] We have been able to discern three possible reasons for this change in the morning results from the results of B08. (1) The primary reason appears to be changing the averaging area from area B to area B′. We know that the statistics of storm behavior depend on location. The diurnal cycle of rainfall in the 2.5° × 2.5° grid box in the southeast corner of area B′ peaks quite strongly in the middle of the afternoon (~1500–1900 LT), whereas the diurnal cycle in the southwest corner of area B is ill defined [e.g., *Hirose et al.*, 2008]. This suggests that the conditions for storm invigoration may be very different in the mornings for the two grid boxes, perhaps enough to change the overall statistics. The other two reasons for the changes in statistics seem to be (2) that the histograms for each hour of the day and day of the week were smoothed by B08 over a 4–5-h range before computing the ratios, whereas they are not smoothed here, and (3) that local times were computed for each 2.5° × 2.5° grid box by B08, whereas the local times at the center of area B′ used here are calculated to determine whether data are assigned to mornings or afternoons. The changes in statistics due to these last two reasons were relatively minor compared with the change resulting from the shift from area B to area B′.

5. Discussion and Conclusions

[20] The increase in afternoon storm heights during the middle of the week compared with weekends seen in

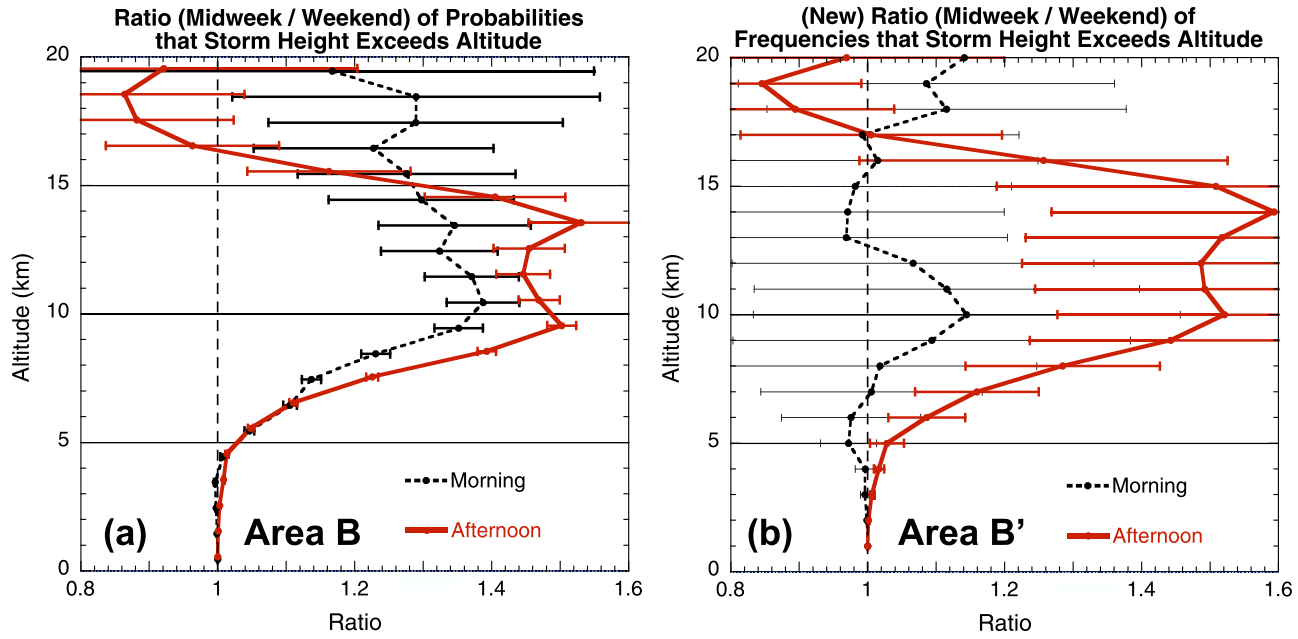

Figure 2. The ratio of the frequency that storm heights exceed a given altitude during the midweek (Tuesday–Thursday) to the frequency that weekend (Saturday–Monday) storm heights exceed that altitude. Dashed black lines are for morning (0000–1200 LT) data; red lines are for afternoon (1200–2400 LT) data. (a) Ratio over area B (figure reproduced from *Bell et al.* [2008]). (b) New results for area B′ with 1-sigma error bars estimated here. A ratio of 1 (thin dashed line) would be expected if there were no variation in storm height distributions with the day of the week.

Figure 2 is consistent with the physical picture that storm growth is enhanced by the presence of additional particulate pollution in the atmosphere during the middle of the week. The size of the increase appears to be statistically strongest at higher altitudes, and it would be hard to attribute the increases to the happenstances of sampling.

[21] The result for morning storm heights obtained here, that there is no clear change in behavior with the day of the week, is easier to understand than the behavior found by B08 for morning storms: the physical mechanism invoked to explain the afternoon changes would suggest that since convective potential is smaller in the morning hours and there is likely to be less release of latent heat of fusion because of the freezing of water droplets, there should be less invigoration of morning storms and less of a weekly cycle. The sensitivity of the morning results to the exact area over which the statistics are obtained, possibly owing to the changes in the diurnal variations present in area B, however, suggests that unraveling how pollution affects morning convection will be more difficult.

Appendix A

[22] The sampling error variance estimates for the ratios $r_i(a)$ defined in equation (2) are made using the weekly values of the storm height distributions $n_{p,w}(a)$. We present here some details about how these estimates are made. An expression for the error variance of $r_i(a)$ is given in equation (4) in terms of the error variances of $\delta c_{\mathrm{TWT},i}$ and $\delta c_{\mathrm{SSM},i}$, and we show here how these two variances can be estimated. As in equation (5), we use p to symbolize the period from which the data came, whether TWT or SSM and whether i is m or a (morning or afternoon).

[23] We first define the cumulative sum of the storm height histograms for each week w,

$$N_{p,w}(a) = \sum_{a'=a}^{t} n_{p,w}(a'), \tag{A1}$$

and the cumulative sum of the overall storm height histograms,

$$N_p(a) = \sum_{a'=a}^{t} n_p(a') \tag{A2}$$

$$N_p(a) = \sum_{w=1}^{W} N_{p,w}(a), \tag{A3}$$

which gives the number of PR-observed storm heights in area B′ falling in bin a or higher. The fractions $c_p(a)$ defined in equation (1) can then be written

$$c_p(a) = \frac{N_p(a)}{N_p(0)}. \tag{A4}$$

[24] As was done in equation (3), we describe the sampling error in terms of deviations from the climatological mean,

$$N_p(a) = \langle N_p(a) \rangle + \delta N_p(a). \tag{A5}$$

The first term on the right-hand side of (A5) is the expected count, and the second term is the deviation from the expected count due to the particular data sample from which

we happened to have calculated $N_p(a)$ from. We can use analogous notation to rewrite equation (A4) as

$$c_p(a) = \langle c_p(a)\rangle + \delta c_p(a) \tag{A6}$$

$$c_p(a) = \frac{\langle N_p(a)\rangle + \delta N_p(a)}{\langle N_p(0)\rangle + \delta N_p(0)}. \tag{A7}$$

[25] Assuming that the fluctuations $\delta N_p(0)$ are not too big relative to $\langle N_p(0)\rangle$, we expand equation (A7) to second order in δ and then keep only second-order terms in δ for the expression for $\langle[\delta c_p(a)]^2\rangle$ to obtain

$$\langle[\delta c_p(a)]^2\rangle \approx \frac{\langle[\delta N_p(a)]^2\rangle}{\langle N_p(0)\rangle^2} - 2\frac{\langle N_p(a)\rangle}{\langle N_p(0)\rangle}\frac{\langle \delta N_p(a)\delta N_p(0)\rangle}{\langle N_p(0)\rangle^2} + \frac{\langle N_p(a)\rangle^2}{\langle N_p(0)\rangle^2}\frac{\langle[\delta N_p(0)]^2\rangle}{\langle N_p(0)\rangle^2} \tag{A8}$$

$$\langle[\delta c_p(a)]^2\rangle = \frac{1}{\langle N_p(0)\rangle^2}\left\{\langle[\delta N_p(a)]^2\rangle - 2c_p(a)\langle\delta N_p(a)\delta N_p(0)\rangle + c_p^2(a)\langle[\delta N_p(0)]^2\rangle\right\}. \tag{A9}$$

Following the approach that led to equation (9), we can estimate the first and last terms inside the braces in equation (A9), using the fact that $N_p(a)$ can be written as a sum of weekly contributions (equation (A3)). We find

$$\langle[\delta N_p(a)]^2\rangle \approx W\mathrm{Var}[N_{p,w}(a)], \tag{A10}$$

where the variance on the right-hand side is calculated as in equation (8).

[26] The same approach allows us to estimate the middle term inside the braces of equation (A9). One finds

$$\langle\delta N_p(a)\delta N_p(0)\rangle \approx W\mathrm{Cov}[N_{p,w}(a), N_{p,w}(0)] \tag{A11}$$

with the covariance estimated in the usual way: letting $x(w) = N_{p,w}(a)$ and $y(w) = N_{p,w}(0)$,

$$\mathrm{Cov}[x(w), y(w)] \approx \frac{1}{W-1}\cdot\sum_w [x(w)-\overline{x}][y(w)-\overline{y}]. \tag{A12}$$

[27] Given equation (4) and the estimates of $\mathrm{Var}[\delta c_p(a)]$ provided by equations (A9)–(A11) here, we can estimate the error variance for the ratios r_i. The square roots of these error variances are used as 1-sigma error bars in Figure 2.

[28] Note that because morning and afternoon samples may only be separated by a few hours, sampling errors for the morning and afternoon cases may not be statistically independent of each other. Consequently, the error bars in Figure 2 may not be appropriate for testing whether the ratios r_m and r_a are statistically different from each other. Sampling errors in $r_i(a)$ at one altitude are almost certainly correlated with those at other altitudes as well.

[29] **Acknowledgments.** Research by T.L.B. was supported by the Science Mission Directorate of the National Aeronautics and Space Administration as part of the Precipitation Measurement Mission program under Ramesh Kakar.

References

Andreae, M. O., D. Rosenfeld, P. Artaxo, A. A. Costa, G. P. Frank, K. M. Longo, and M. A. F. Silva-Dias (2004), Smoking rain clouds over the Amazon, *Science*, *303*(5662), 1337–1342.

Bell, T. L., D. Rosenfeld, K.-M. Kim, J.-M. Yoo, M.-I. Lee, and M. Hahnenberger (2008), Midweek increase in U.S. summer rain and storm heights suggests air pollution invigorates rainstorms, *J. Geophys. Res.*, *113*, D02209, doi:10.1029/2007JD008623.

Hirose, M., R. Oki, S. Shimizu, M. Kachi, and T. Higashiuwatoko (2008), Finescale diurnal rainfall statistics refined from eight years of TRMM PR data, *J. Appl. Meteorol. Climatol.*, *47*, 544–561.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), NCEP-DOE AMIP-II reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*, 1632–1643, doi:10.1175/BAMS-83-11-1631.

Negri, A. J., T. L. Bell, and L. Xu (2002), Sampling of the diurnal cycle of precipitation using TRMM, *J. Atmos. Oceanic Technol.*, *19*, 1333–1344.

Remer, L. A., et al. (2005), The MODIS aerosol algorithm, products and validation, *J. Atmos. Sci.*, *62*, 947–973.

Rosenfeld, D., U. Lohmann, G. B. Raga, C. D. O'Dowd, M. Kulmala, S. Fuzzi, A. Reissell, and M. O. Andreae (2008), Flood or drought: How do aerosols affect precipitation?, *Science*, *321*(5894), 1309–1313, doi:10.1126/science.1160606.

Williams, E., et al. (2002), Contrasting convective regimes over the Amazon: Implications for cloud electrification, *J. Geophys. Res.*, *107*(D20), 8082, doi:10.1029/2001JD000380.

T. L. Bell, Laboratory for Atmospheres, NASA Goddard Space Flight Center, Mail Code 613.2, Greenbelt, MD 20771, USA. (thomas.l.bell@nasa.gov)

M.-I. Lee, NASA Goddard Space Flight Center, Mail Code 610.1, Greenbelt, MD 20771, USA. (myong-in.lee-1@nasa.gov)

J.-M. Yoo, Department of Science Education, Ewha Womans University, 11-1, Daehyun-Dong, Seodaemun-Gu, Seoul 120-750, South Korea. (yjm@ewha.ac.kr)